Assessing the global influence of ENSO on flood risk through 1600 years of simulations

Lenin Del Rio Amador[1], Mathieu Boudreault[1], and David A. Carozza[1]

[1]Université du Québec à Montréal

December 7, 2022

Abstract

El Niño-Southern Oscillation (ENSO) is often considered as a source of long-term predictability for extreme events via its teleconnection patterns. However, given that its characteristic cycle varies from two to seven years, it is difficult to obtain statistically significant conclusions based on observational periods spanning only a few decades. To overcome this, we apply the global flood risk modeling framework developed by Carozza and Boudreault to an equivalent of 1600 years of bias-corrected GCM outputs. The results show substantial anomalies in flood occurrences and impacts for El Nino and La Nina when compared to the all-year baseline. We were able to obtain a larger global coverage of statistically significant results than previous studies limited to observational data. Asymmetries in anomalies for both ENSO phases show a larger global influence of El Nino than La Nina on flood hazard and risk.

Assessing the global influence of ENSO on flood risk through 1600 years of simulations

L. Del Rio Amador[1], **M. Boudreault**[1], **D. A. Carozza**[1,2]

[1]Département de Mathématiques, Université du Québec à Montréal, Montréal, QC, Canada
[2]Currently at Institutional Shareholder Services Canada Inc., Toronto, ON, Canada

Key Points:

- We simulated an equivalent of 1600 years of realistic flood events globally using a statistical model forced with climate model outputs.
- We found that ENSO has statistically significant impacts on a larger share of basins than what was previously found with observational data.
- Asymmetries in anomalies for both ENSO phases show a larger global influence of El Niño than La Niña on flood hazard and risk.

Corresponding author: L. Del Rio Amador, lenindelrio@gmail.com

Abstract

El Niño-Southern Oscillation (ENSO) is often considered as a source of long-term predictability for extreme events via its teleconnection patterns. However, given that its characteristic cycle varies from two to seven years, it is difficult to obtain statistically significant conclusions based on observational periods spanning only a few decades. To overcome this, we apply the global flood risk modeling framework developed by Carozza and Boudreault to an equivalent of 1600 years of bias-corrected GCM outputs. The results show substantial anomalies in flood occurrences and impacts for El Niño and La Niña when compared to the all-year baseline. We were able to obtain a larger global coverage of statistically significant results than previous studies limited to observational data. Asymmetries in anomalies for both ENSO phases show a larger global influence of El Niño than La Niña on flood hazard and risk.

Plain Language Summary

Global assessment of the occurrence probability and impact of floods is of key interest to environmental research, climate science, economics and financial risk management of flooding (governments, insurance and reinsurance industry, banks). However, hydrological models are too complex to evaluate the links between possible harms and risks associated to floods and climate variability at global scales for long periods of time. Only a few studies have been performed in this direction, but they are limited to observational data spanning only a few decades. In this paper, we used the statistical and machine learning modeling framework developed by Carozza and Boudreault to relate flood hazard and risk to El Niño-Southern Oscillation (ENSO), which is the main driver of interannual climate variability and one of the most predictable phenomena at these time scales. By producing an equivalent of 1600 years of simulations consistent with global climate models, we were able to obtain statistically significant results for a larger portion of the planet than previous studies limited to observational data. We also found a greater global influence of El Niño than La Niña on flood hazard and risk.

1 Introduction

Interannual climate variability is dominated by the El Niño-Southern Oscillation (ENSO) signal (H.-J. Wang et al., 1999). Its mechanisms of teleconnections and influence over the climate at global scales have been vastly studied using many different approaches, e.g. dynamical (C. Wang, 2018; Liu & Alexander, 2007; Domeisen et al., 2019), climate networks (Tsonis et al., 2006; Tsonis, 2018; Zhou et al., 2015), stochastic (Del Rio Amador & Lovejoy, 2021a), empirical/statistical (Rashid, 2020; Penland & Sardeshmukh, 1995) and stochastic-dynamical (N. Chen & Majda, 2017; Giorgini et al., 2022). Besides conventional General Circulation Models (GCMs), these models have been applied to obtain skilful predictions of ENSO with lead times up to several months (X. Wang et al., 2020; L'Heureux et al., 2019; Penland & Magorian, 1993).

In contrast to the atmosphere, which exhibits deterministic predictability limits of ≈ 7 to 10 days (the lifetime of planetary-size structures), the corresponding limit for ocean temperatures can go up to 2 years (Lovejoy & Schertzer, 2012, 2013; Lovejoy et al., 2018; Del Rio Amador & Lovejoy, 2021b). This is implicitly evidenced by coupled GCMs, which can predict low-frequency sea surface temperature (SST) variabilities such as ENSO and the Pacific Decadal Oscillation (PDO) at lead times of up to two years (D. Chen et al., 2004; Choi & Son, 2022). There are currently more than 20 models on ENSO for 3-month average real-time forecasts of the next 9 months (IRI, 2022). This makes the ENSO phenomenon the most predictable target of seasonal climate forecast.

The main interest in forecasting ENSO comes from its strong correlation with episodes of rainfall (Shukla & Paolino, 1983), snowfall (Patten et al., 2003), droughts (Yu & Zou, 2013; Kumar et al., 2006), hurricanes (Pielke & Landsea, 1999; Kim et al., 2009; G. Chen

& Tam, 2010; Zhang et al., 2015) and severe temperature patterns (Yang et al., 2018; Ropelewski & Halpert, 1986; Halpert & Ropelewski, 1992; Weng et al., 2009). Preparation for such extreme events is essential for decision makers in order to mitigate their impact (de Perez et al., 2014). However, given that the characteristic cycle of El Niño and La Niña patterns varies from two to seven years, it is difficult to achieve statistically significant conclusions based on observational periods spanning only a few decades. For instance, although ENSO is known to influence hydrology and precipitation patterns in many regions of the world, only a few studies explored its impact on flood risk globally (P. J. Ward, Eisner, et al., 2014; P. J. Ward, Jongman, et al., 2014; P. Ward et al., 2016; Yan et al., 2020; Corringham & Cayan, 2019; Saghafian et al., 2017). All these analyses have been performed using observational time series spanning less than 50 years, i.e. around only a dozen El Niño events. As Emerton et al. (2017) pointed out, "the likelihood of increased or decreased flood hazard during ENSO events is much more complex than is often perceived and reported" due to the limited length and the uncertainties inherent in the data.

In addition, the simulation of global scenarios of flood risk that are consistent from climate, hydrological, hydraulic, and exposure standpoints is also limited by the computational cost of regional hydrological models and the data requirements that are not necessarily available globally (P. J. Ward et al., 2015). Typical approaches usually force higher-resolution hydrological models with runoff from lower-resolution climate model outputs (Winsemius et al., 2013; Yamazaki et al., 2011), adding an extra layer of complexity and uncertainty to the final result. This makes unpractical the use of relatively large series of climate model simulations to force hydrological flood models with the purpose of studying the influence of ENSO on flood risk.

To overcome this lack of sufficiently long global series of observations or hydrological simulations, in this paper we use the global flood modeling framework developed by Carozza and Boudreault (C&B in the following) (Carozza & Boudreault, 2021). This data-driven model is climate-consistent, global, fast, flexible, and is ideal for applications that do not necessarily require high-resolution flood mapping. It applies statistical and machine learning methods to relate historical flood occurrence and impact data with climatic, watershed, and socioeconomic factors for 4,734 basins at Pfafstetter level 5 globally.

The relatively low computational cost of the C&B framework allows us to simulate an equivalent of 1600 years of flood hazard and risk by combining it with bias-corrected outputs from GCMs. Our goal is to obtain statistically robust distributions of the influence of El Niño on flood risk at a global scale. It builds on the capacity of the C&B framework to replicate the actual occurrence and impact of floods from environmental variables and the ability of climate models to reproduce global patterns of ENSO events. This is of key interest to flood and environmental research, climate science, economics and financial risk management of flooding (governments, insurance and reinsurance industry, banks).

2 Data and Methods

In the following paragraphs, we briefly describe the methods and data used in the C&B global flood risk modeling framework to simulate the occurrence and impact of floods around the globe, based on environmental and socioeconomic factors. We then present the GCM ensemble chosen to feed the model to produce long series of flood events which are physically consistent with the climate dynamics. Finally, we discuss the index used to characterize ENSO for each model output, aiming to identify statistical relationships between annual flood hazard and risk and the ENSO cycle.

2.1 Global Flood Risk Modeling Framework

The C&B global flood risk modeling framework, introduced by Carozza and Boudreault (2021), is driven by historical flood and environmental observations. As a statistical model, it is capable of quickly generating large global catalogs of flood events that are physically consistent with climate. In the original paper, the authors considered classical and machine learning methods, such as logistic and linear regression (LR), random forests (RF) (Breiman, 2001), and artificial neural networks (NN) (McCulloch & Pitts, 1943), to solve the statistical problems of classification and regression for flood occurrence and impact, respectively. In the present work, we only take outputs from the RF model, since it consistently showed better predictive skill than LR and is easier to interpret than NN models, while still capturing complex non-linear relationships and interactions between predictors.

To train the model, an observational flood occurrence and impact dataset was built by intersecting data of historical flood events from the Dartmouth Flood Observatory Global Active Archive of Large Flood Events (DFO) (Brakenridge, 2019) with the HydroBASINS dataset (Lehner & Grill, 2013) of global watersheds at Pfafstetter level 5. For each of these 4,734 basins, covering the entire global land surface except Antarctica, the model associates annual flood occurrence and impact to the driving climatic, watershed, and socioeconomic factors. In the DFO dataset, each flood is characterized by impact metrics such as the duration, deaths caused, population displaced, and severity (a proxy of return period). Here, we choose the population displaced as a measure of impact. This could then be translated into a measure of economic impact by simply multiplying population displaced by the annual gross domestic product (GDP) per capita based on purchasing power parity (PPP) (Kummu et al., 2018) of a given watershed. This "GDP disrupted" does not directly measure all the economic losses, but could be regarded as a proxy of the economic impact associated to a flood event.

For the case of flood occurrence, a total of 38 predictors were considered: average temperature for the hydrological year (October 1–September 30) and hydrological annual maximum precipitation at four different timescales were used as climate predictors; 31 time-invariant covariates were taken to represent watershed characteristics, location and storage capacity; and finally, population density and GDP per capita were used as time-varying proxies of urbanization and flood control. To model the flood impact, the same predictors were used, only replacing the annualized values for temperature and precipitation by average values over 7, 8–30, 31–60, and 61–120 days, prior to each flood event, for a total of 41 independent covariates.

The historical data for precipitation was built by combining the Climate Hazards Group Infrared Precipitation with Stations (CHIRPS) dataset (Funk et al., 2015) for latitudes from 50°S to 50°N, and the CPC Global Unified Gauge-Based Analysis of Daily Precipitation (CPC Precipitation) dataset (Xie et al., 2007) for all other latitudes. The temperature was taken from the CPC Global Daily Temperature (CPC Temperature) dataset (Shi, 2007). A full description and references to the data used for all predictors are provided in Carozza and Boudreault (2021).

The model was validated with observations from the DFO database in the period 1985–2017 for a total of 32 hydrological years. Considering the 4734 watersheds at Pfafstetter level 5, there are 151,488 occurrence observations that can be either "flood" or "no flood". After removing observations with missing data in the predictors, we are left with 128,494 observations for fitting the occurrence model, of which 19,746 are positive events used to fit the impact component. Out-of-sample cross-validation was always performed using a random sampling of 70% for training and the remaining 30% as a test set. Carozza and Boudreault (2021) report competitive values of skill score metrics for both components of the model, reflecting the ability of the C&B framework to predict flood hazard and impact over most of the globe.

To further confirm the quality of the model and as an example application, Carozza and Boudreault (2021) stochastically simulate 1 million years of flood occurrences and impacts over 4,734 watersheds globally. This is achieved by replacing the time-varying climate predictors by bias-corrected outputs from the National Center for Atmospheric Research's (NCAR) Community Earth System Model (CESM) Large Ensemble (LE) (Kay et al., 2015). Using simulated temperature and precipitation data from the 40 members of CESM-LE in the 40-year period 1980–2020 (consistent with the flood observational record from DFO), the authors were able to obtain physically consistent flood hazard and risk distributions for an equivalent of 1600 years. The 1 million years of events were produced by sampling from these distributions. The good agreement between the simulated and observational values of flood occurrence and impact is another validation of the quality of the C&B framework.

2.2 Climate Model Output Data

Our goal in this paper is to study the influence of ENSO on flood hazard and risk from the simulations produced by Carozza and Boudreault (2021). Besides the possibility to obtain long series of flood events by combining independent outputs from its 40 members, we also choose the CESM-LE for its highlighted skill on reproducing ENSO events and its associated teleconnetion patterns (Deser et al., 2012; Vega-Westhoff & Sriver, 2017). The CESM (Hurrell et al., 2013) is a fully coupled climate model composed of seven modules: atmosphere, land, river runoff, ocean, sea ice, land ice, and ocean wave. It was introduced by NCAR to examine interannual climate variability in the context of anthropogenic climate change and focused on improving the modeling of ENSO features, including its asymmetry and diversity, by introducing the westerly wind bursts parameterization (Tan et al., 2020).

The CESM-LE is a set of 40 independent runs simulating the Earth system for the years 1920–2100 that share the same forcing of radiative gases in the atmosphere and aerosols. To achieve independence, each member is initialized with a roundoff error perturbation to the atmosphere in model year 1850. Here we only use the outputs for precipitation (rainfall + snowfall) from the Community Land Model 2.0 (Lawrence et al., 2011), as well as the temperature 2m above the surface from the Community Atmosphere Model 5.2 (Neale & Group, 2012). We also limited the series to the period 1980–2020 to match the flood observational record from DFO used to fit the statistical model. We did so because of the known inability of the RF algorithm to extrapolate out of the training domain (Hastie et al., 2009).

Similar to how it was done with the climate predictors used in the statistical fit step, the precipitation and temperature from CESM-LE (originally at 0.5° resolution) were aggregated by averaging over the grid points in each level 5 watershed. The methodology of Hempel et al. (2013) was applied to these aggregated values to correct for biases in precipitation relative to CHIRPS and CPC Precipitation and in temperature relative to CPC Temperature. This debiasing method to correct monthly means and daily variability about the means is widely used in the hydrological and flood impact literature.

2.3 ENSO Index

As mentioned earlier, we choose the CESM-LE for its ability to replicate global patterns associated to ENSO events. To study the influence of the latter on flood risk, we applied the C&B model to simulate flood hazard and impact using precipitation and temperature for each member of the CESM-LE outputs, while sea surface temperature (SST) from the same ensemble member was used to obtain the corresponding ENSO index. In this way, the physical links that relate ENSO to flood occurrence and intensity are preserved through the internal dynamics of the climate model.

There are many indices that are typically used to define the phase and strength of ENSO events. They include regional SST–based indices [e.g.: Niño-1+2, Niño-3, Niño-4, Niño-3.4, and Japan Meteorological Agency (JMA)], the surface atmospheric pressure–based Southern Oscillation index (SOI), and other more complex definitions such as the trans-

Niño index (TNI) and the multivariate ENSO index (MEI). Each of them has its own benefits and disadvantages. A detailed description and intercomparison among the indices mentioned above is given by Hanley et al. (2003). In the present study, we choose the JMA index because of its good sensitivity on selecting ENSO events (Bove et al., 1998).

The Japan Meteorological Agency defines the ENSO index as a 5-month running mean of spatially averaged SST anomalies over the tropical Pacific: 4°S–4°N, 150°W–90°W (similar to the region used for the Niño-3, but with 1° reduction in latitude). To identify ENSO years, they use the same definition that we use to identify annual flood events for hydrological years: from October through the following September. If the index value is above 0.5°C for at least 6 consecutive months (including October-November-December), the ENSO year is categorized as El Niño, if it is below -0.5°C as La Niña, and as neutral for all other values. Another advantage of the JMA index is that the ± 0.5°C thresholds used to determine the different phases is very close to the more general range defined from the 25%–75% quantiles: -0.52°C to 0.47°C for JMA index in the period 1894–1993 (Hanley et al., 2003). It is also worth mentioning that the correlation between the ENSO-JMA index and other ENSO indices is very high (above 0.9) and as such, we feel that using another index to identify ENSO phases would not influence our conclusions.

3 Results

In this section, we first analyze the relationship between annual flood occurrence and ENSO without distinguishing individual phases. Then, we present global maps of flood impact anomalies separating El Niño and La Niña events. This distinction allows us to detect asymmetries in the influence of ENSO on flood risk. We compare our results with previous studies that perform similar analyses at a global scale or for specific regions.

3.1 ENSO Influence on Flood Occurrence

To simulate the annual occurrence of floods from the CESM-LE outputs, the C&B model performs Bernoulli trials using the probability parameter obtained from the RF algorithm. The RF was fitted with observational data and forced with bias-corrected precipitation and temperature for each CESM ensemble member year and watershed (40 members $\times$ 40 years $\times$ 4,734 watersheds). In Fig. 1a, we show an example time series of flood occurrence probability (black curve) from one of the ensemble members for a random basin centered at 14.7°N 85.8°W (in Honduras). From the same GCM output, we used SST to compute the annually averaged ENSO JMA index, shown as coloured bars in Fig. 1a. The ENSO phases for each hydrological year were identified according to the criteria described in Section 2.3.

Similar series were obtained for each of the 40 CESM-LE outputs, giving a total of 1600 points to compute the Pearson correlation between the occurrence probability and the annually averaged ENSO JMA index for each basin. The results are shown in Fig. 1b, where watersheds without enough observational data to perform the fit are shown in black, and locations where the correlation is different from zero with less than 95% confidence are shown in white (absolute value is less than 0.05, considering 1600 pairs of data). The black area (corresponding to high latitudes lacking topographical observations for the predictors) represents approximately 12% of the global surface over land (excluding Antarctica).

The correlation patterns shown in Fig. 1b generally agree with regional results reported for Asia (Iqbal & Hassan, 2018; Saghafian et al., 2017), North America (Corringham & Cayan, 2019; Hamlet & Lettenmaier, 2007), South America (Isla & Junior, 2013), Australia (Kiem et al., 2003), Europe (Nobre et al., 2017) and Africa (Nicholson & Kim, 1997). To our knowledge, there are only a few studies in the literature reporting links between climate oscillations and floods at global scales [see the review by Kundzewicz et al. (2019)], most of them led by the VU Amsterdam group (P. J. Ward, Eisner, et al., 2014; P. J. Ward,

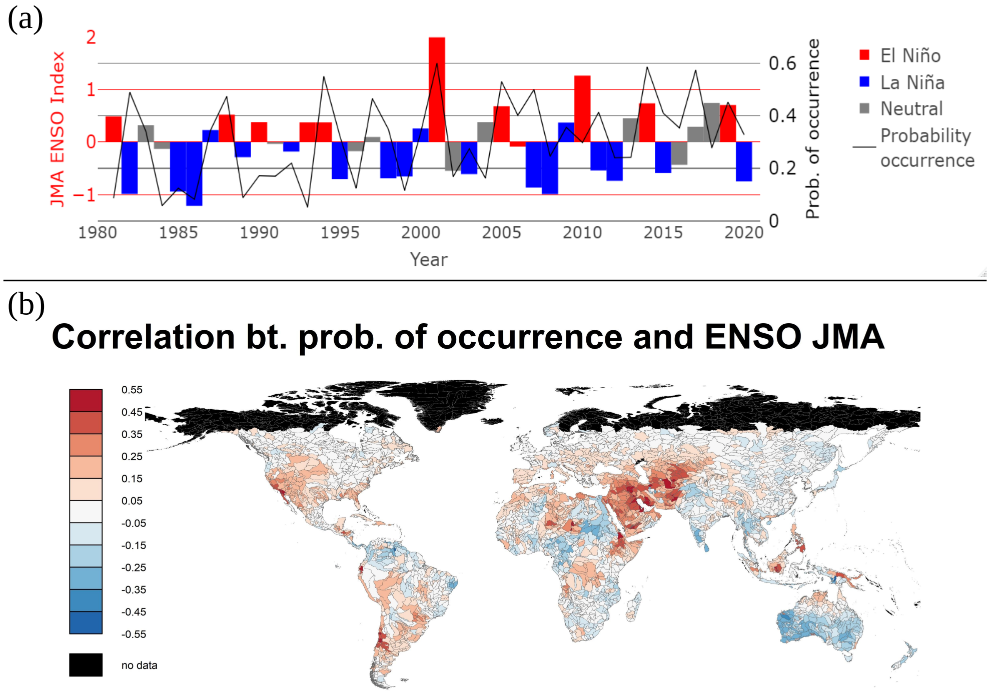

Figure 1. a) Example time series of occurrence probability (in black) from one of the ensemble members for a basin centered at 14.7°N 85.8°W, and the corresponding ENSO index for the same GCM output. b) Pearson correlations between the combined 1600-point time series of occurrence probabilities and the annually averaged ENSO JMA index. Locations with correlation different from zero with less than 95% confidence are shown in white. Watersheds without enough data to perform the fit are shown in black.

Jongman, et al., 2014; P. Ward et al., 2016; Yan et al., 2020). In general, the correlation patterns shown in Fig. 1b match those reported by Ward et al. with a few exceptions.

In P. J. Ward, Eisner, et al. (2014), the authors report significant correlation (with more than 90% confidence) for 37% of land area. In the map shown in Fig.1b (despite the more restrictive criterion of zero being out of the 95% confidence interval) we get significant correlation for 55% of the total land area (complementing the 33% not significant plus 12% corresponding to no data). This makes the present work the study with the largest global coverage of significant correlation between flood hazard and ENSO to date.

From the significant correlation values, 35% of the overall land area are positive (higher occurrence probability during El Niño and lower during La Niña) and 20% of the total land surface are negative (lower during El Niño and higher during La Niña). This contradicts the results from P. J. Ward, Eisner, et al. (2014), where they find larger land surface with negative correlation (23%) than with positive significant correlation (14%). Notice that in their paper the signs are inverted, since they use the Southern Oscillation Index (SOI) to characterize ENSO, which is in opposite phase with respect to the JMA index used here. It is worth mentioning that they only use 41 years of observations with 10 El Niño and La Niña events.

3.2 ENSO Influence on Flood Impact

The correlations reported in the previous section provide an overall idea of the influence of ENSO on flood occurrence at the global scale. However, they do not allow the individual effects of each of the phases to be distinguished. In the following, we make this distinction to analyze the relationship between ENSO and flood risk, measured through population displaced and GDP disrupted.

In Fig. 2a and b, we show maps of anomalies averaged over El Niño and La Niña phases, respectively. These variations are taken with respect to the mean population displaced by flood events considering all the hydrological years [see Fig. 10 in (Carozza & Boudreault, 2021)]. Regions where anomalies are different from zero with less than 95% confidence are shown in white, while locations where not enough data were available for the analysis are shown in black. In contrast with other studies that only consider limited periods of observational data, in these figures we are able to clearly and reliably identify opposite global patterns of anomalies for the positive and the negative phases of ENSO.

The opposite symmetry for the warm and the cold phases is expected by definition: since there are almost the same number of El Niño than La Niña years, any possible imbalance in the anomalies with respect to the all-year baseline could only be attributed to the Neutral phase. This symmetry is broken in a few locations due to the intrinsic nonlinear dynamics of the climate system. For example, the impact of flooding over Japan and Sri Lanka is large in La Niña years, while no significant deviation from the average was detected during El Niño. Anomalies with the same sign were also detected for the two phases, e.g.: in Ecuador (negative-negative) and Belgium (positive-positive).

While population displaced by flooding is one proxy of its impact, a measure like GDP disrupted as defined above should be a more direct measure of economic losses. The corresponding maps of anomalies in GDP disrupted are shown in Fig. 2c and d. Although the patterns observed during El Niño or La Niña phases are generally similar to those of population displaced, the GDP maps give direct information relevant for policymakers and for risk analysis in the financial, insurance and reinsurance industries.

The differences between the population displaced and the GDP disrupted patterns arise from the heterogeneous global distribution of wealth. For instance, a direct comparison between the top and the bottom panels of Fig. 2, shows that a similar number of people displaced relative to the mean translates into much higher economic losses anomalies in the United States and Europe than in Central Africa, i.e. the colours intensify for the first and gets lighter for the latter when you switch from one measure to the other. In South America and Asia, the colouring remains similar for the two impact measures, showing intermediate values of GDP per capita with respect to the regions mentioned before. However, the combined effect of flood intensity and economic exposure is not trivial, since high-income countries could have overpopulation in affected urban areas while low-income countries could present higher vulnerability because of lower investments in risk reduction measures, among other causes.

The information presented in the previous maps is very useful since it provides an overall idea of flood risk in terms of population displaced or GDP disrupted. But it would also be interesting to better differentiate flood *hazard* from flood *risk*. As such, we computed a unit-less metric by normalizing the anomalies with respect to the all-year impact. The results are shown in Fig. 3, where the average difference of number of people displaced during (a) El Niño and (b) La Niña years is expressed as a percentage of the average number of people displaced considering all years (very similar maps are obtained if the GDP disrupted is used instead). Such measure is thus closer to represent flood hazard in terms of flood occurrence and intensity.

There are clear differences between the patterns shown in Fig. 3, and the corresponding maps in Fig. 2, a and b. For example, over India and China there are large opposite values of anomalies reported for El Niño and La Niña phases, but they only represent a small percentage of the average number of people affected by floods in these regions. On the contrary, in Australia, the difference in number of people displaced for each

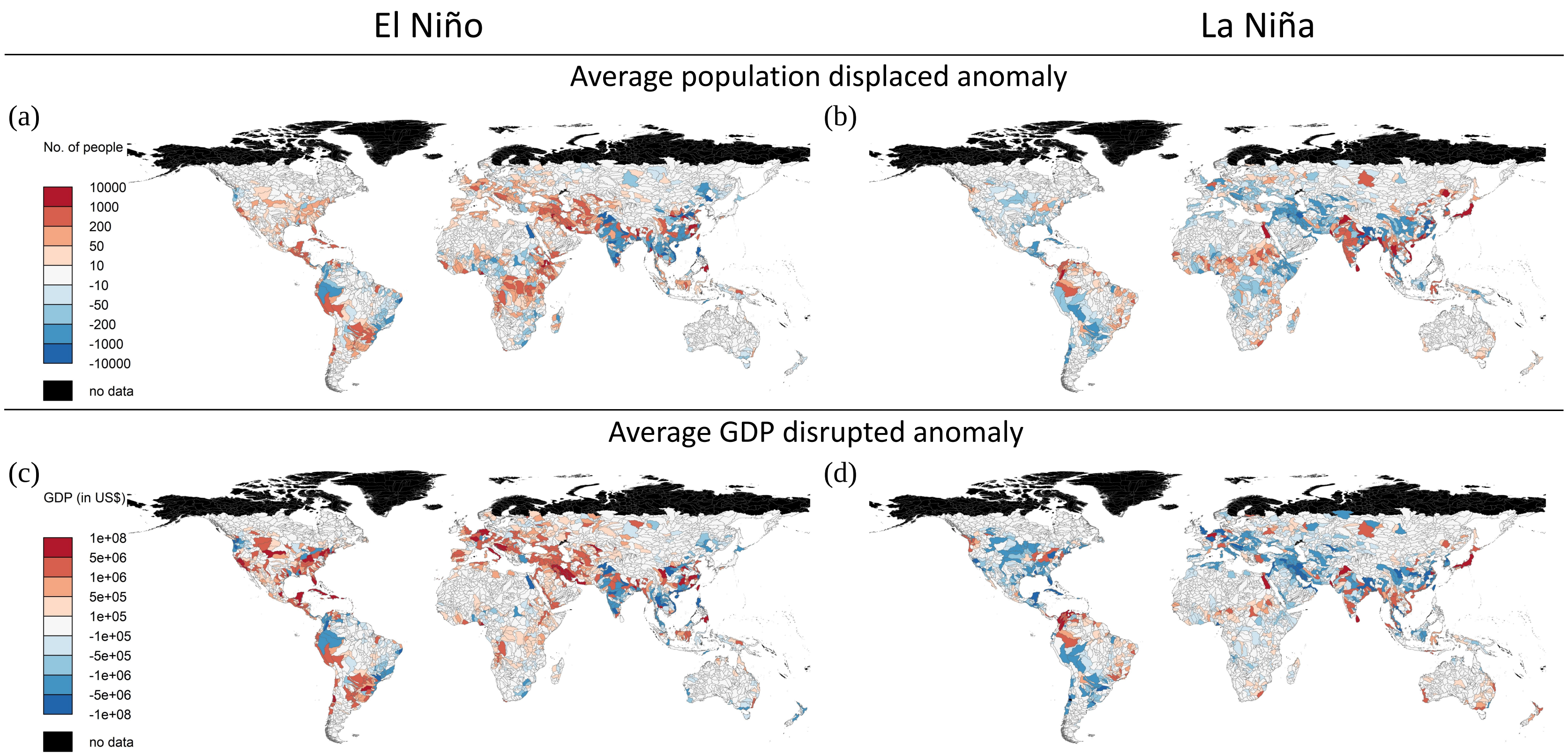

Figure 2. Anomalies in population displaced (top) and GDP disrupted (bottom) with respect to the all-year baseline for (a,c) El Niño and (b,d) La Niña phases. Regions where anomalies are different from zero with less than 95% confidence are shown in white. Basins in black correspond to not enough data for the fit.

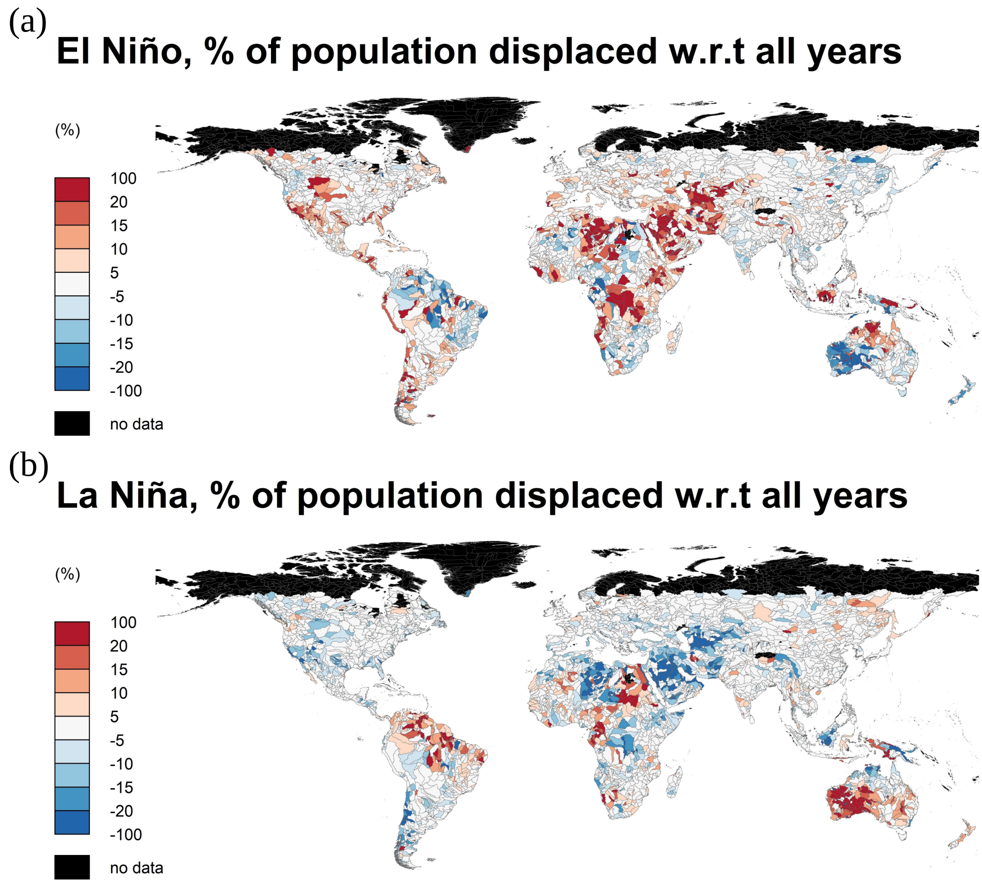

Figure 3. Anomalies of the average number of people displaced during (a) El Niño and (b) La Niña years as a percentage of the average number of people displaced considering all years. Regions where the anomalies are different from zero with less than 95% confidence are shown in white. Basins in black correspond to not enough data for the fit.

phase is relatively small, but the normalized anomalies are significantly large since the average impact for all years is also low, except for the Eastern part of the country [see Fig. 10 in (Carozza & Boudreault, 2021)]. In this region, although it is considerably affected by floods, the impact seems to be uniformly distributed over all ENSO phases.

We also remark that the maps shown in Fig. 3 very well replicate the reported global patterns of ENSO-induced precipitation (Dai & Wigley, 2000). They also agree relatively well with global maps of flood risk obtained by P. J. Ward, Eisner, et al. (2014); P. J. Ward, Jongman, et al. (2014); Yan et al. (2020). However, as we mentioned earlier, we obtain larger global coverage of statistically significant anomalies as well as different results regarding the asymmetric global influence of ENSO on flood impact. The normalized anomalies for both phases are significant with more than 95% confidence for 69% of the global land surface (excluding Antarctica). For the La Niña phase, we find that: 0.6% less people are affected globally with respect to the overall average (all years), 0.2% less of the global GDP is disrupted, 29% of the total surface affected corresponds to significant positive anomalies, while 40% is negative. For El Niño, we have: 0.4% more people and 0.2% more of the global GDP are affected, the surface partition is 41% positive while 28% has negative anomalies. This is in agreement with the correlation asymmetry mentioned in Section 3.1. In general, El Niño shows a greater global impact on flood hazard and risk than La Niña.

4 Summary and Conclusions

ENSO is one of the main and most predictable components of interannual climate variability. A large amount of evidence of its relation with the frequency and intensity of extreme events has been published. However, studies showing its global influence directly on flood occurrence and impact have been limited by the lack of sufficiently long global series of observations (comprising only a few ENSO cycles), and by the high computational cost of hydrological models to obtain long series of simulations.

In this paper, we used the empirical C&B global flood risk modeling framework to simulate an equivalent of 1600 years of realistic flood events for each of 4,734 basins globally. The simulations were created by forcing the statistical model with bias-corrected precipitation and temperature output from the large ensemble of the NCAR CESM climate model. SST outputs from the same GCM were used to obtain ENSO indices for the same 1600 hydrological years. This approach allowed us to obtain physically consistent relationships between floods and ENSO with high a degree of confidence from a statistical point of view. Our results rely on the ability of the C&B framework to replicate the actual occurrence and impact of floods and the skill of the climate model to replicate ENSO events.

The maps presented show similar distributions of normalized flood impact anomalies than known global patterns of ENSO-induced precipitation. They also agree relatively well with previous studies on flood risk at both global and regional scales, but we identified some important discrepancies. The much longer simulation periods used in our study allowed us to observe more frequent opposite patterns in flood risk for the warm and the cold phases of ENSO. Observing such patterns was difficult in other studies due to the internal variability and the lack of enough data to improve the signal-to-noise ratio. For the same reason, we were able to obtain reliable values of anomalies in many more regions than previous publications limited to observational data. To our knowledge, the results presented here have the largest global coverage of statistically significant correlations between flood hazard and ENSO to date. The same applies to the significance of the impact anomalies corresponding to each phase. Besides the expected opposite patterns of anomalies for both ENSO phases, we found a symmetry breaking in some regions, with El Niño showing greater global impact than La Niña on flood hazard and risk, in contradiction with P. J. Ward, Eisner, et al. (2014).

5 Data Availability Statement

Datasets, fitted statistical models, simulated catalogs and software for this research are available at `https://doi.org/10.5281/zenodo.3873422` and Carozza and Boudreault (2021). The CESM Large Ensemble dataset is available at `https://www.cesm.ucar.edu/projects/community-projects/LENS/data-sets.html` and the authors acknowledge CESM Large Ensemble Community Project.

Acknowledgments

This work was supported by Mitacs through the Mitacs Accelerate program. This work was partially funded by AXA XL, the property & casualty and specialty risk division of AXA. We acknowledge the support of the Fonds de recherche du Québec–Nature et technologies (FRQNT) [funding references number 119908 and 318522]. Cette recherche a été financée par le Fonds de recherche du Québec–Nature et technologies (FRQNT) [numéros de référence 119908 et 318522]. We acknowledge the support of the Natural Sciences and Engineering Research Council of Canada (NSERC) [funding reference number PDF-502939-2017]. Cette recherche a été financée par le Conseil de recherches en sciences naturelles et en génie du Canada (CRSNG) [numéro de référence PDF-502939-2017]. This work was also supported by a Marine Environmental Observation, Prediction and Response Network (MEOPAR) Postdoctoral Fellowship Award (PDF-13-2019).

At the time of publication, David A. Carozza is employed by Institutional Shareholder Services Canada Inc., a wholly-owned subsidiary of Institutional Shareholder Services Inc. ("ISS"). This paper presents the authors' opinions and not those of ISS, its clients, affiliates, or employees.

References

Bove, M. C., O'Brien, J. J., Eisner, J. B., Landsea, C. W., & Niu, X. (1998, 11). Effect of el niño on u.s. landfalling hurricanes, revisited. *Bulletin of the American Meteorological Society*, *79*, 2477-2482. doi: 10.1175/1520-0477(1998)079⟨2477:EOENOO⟩2.0.CO;2

Brakenridge, G. R. (2019). *Global active archive of large flood events. dartmouth flood observatory, university of colorado, usa.* Retrieved 2019-10-15, from `http://floodobservatory.colorado.edu/Archives/`

Breiman, L. (2001). Random forests. *Machine Learning*, *45*, 5-32. doi: 10.1023/A:1010933404324

Carozza, D. A., & Boudreault, M. (2021, 4). A global flood risk modeling framework built with climate models and machine learning. *Journal of Advances in Modeling Earth Systems*, *13*. doi: 10.1029/2020MS002221

Chen, D., Cane, M. A., Kaplan, A., Zebiak, S. E., & Huang, D. (2004, 4). Predictability of el niño over the past 148 years. *Nature*, *428*, 733-736. doi: 10.1038/nature02439

Chen, G., & Tam, C.-Y. (2010, 1). Different impacts of two kinds of pacific ocean warming on tropical cyclone frequency over the western north pacific. *Geophysical Research Letters*, *37*, n/a-n/a. doi: 10.1029/2009GL041708

Chen, N., & Majda, A. J. (2017, 2). Simple stochastic dynamical models capturing the statistical diversity of el niño southern oscillation. *Proceedings of the National Academy of Sciences*, *114*, 1468-1473. doi: 10.1073/pnas.1620766114

Choi, J., & Son, S.-W. (2022, 12). Seasonal-to-decadal prediction of el niño–southern oscillation and pacific decadal oscillation. *npj Climate and Atmospheric Science*, *5*, 29. doi: 10.1038/s41612-022-00251-9

Corringham, T. W., & Cayan, D. R. (2019, 7). The effect of el niño on flood damages in the western united states. *Weather, Climate, and Society*, *11*, 489-504. doi: 10.1175/WCAS-D-18-0071.1

Dai, A., & Wigley, T. M. L. (2000, 5). Global patterns of enso-induced precipitation. *Geophysical Research Letters*, *27*, 1283-1286. doi: 10.1029/1999GL011140

Del Rio Amador, L., & Lovejoy, S. (2021a). Long-range forecasting as a past value problem: Untangling correlations and causality with scaling. *Geophysical Research Letters*, *48*. doi: 10.1029/2020GL092147

Del Rio Amador, L., & Lovejoy, S. (2021b). Using regional scaling for temperature forecasts with the stochastic seasonal to interannual prediction system (stocsips). *Climate Dynamics*, *57*, 727-756. doi: 10.1007/s00382-021-05737-5

de Perez, E. C., Monasso, F., van Aalst, M., & Suarez, P. (2014, 2). Science to prevent disasters. *Nature Geoscience*, *7*, 78-79. doi: 10.1038/ngeo2081

Deser, C., Phillips, A. S., Tomas, R. A., Okumura, Y. M., Alexander, M. A., Capotondi, A., ... Ohba, M. (2012, 4). Enso and pacific decadal variability in the community climate system model version 4. *Journal of Climate*, *25*, 2622-2651. doi: 10.1175/JCLI-D-11-00301.1

Domeisen, D. I., Garfinkel, C. I., & Butler, A. H. (2019, 3). The teleconnection of el niño southern oscillation to the stratosphere. *Reviews of Geophysics*, *57*, 5-47. doi: 10.1029/2018RG000596

Emerton, R., Cloke, H. L., Stephens, E. M., Zsoter, E., Woolnough, S. J., & Pappenberger, F. (2017, 4). Complex picture for likelihood of enso-driven flood hazard. *Nature Communications*, *8*, 14796. doi: 10.1038/ncomms14796

Funk, C., Peterson, P., Landsfeld, M., Pedreros, D., Verdin, J., Shukla, S., ... Michaelsen, J. (2015, 12). The climate hazards infrared precipitation with stations—a new environmental record for monitoring extremes. *Scientific Data*, *2*, 150066. doi: 10.1038/sdata.2015.66

Giorgini, L. T., Moon, W., Chen, N., & Wettlaufer, J. S. (2022, 6). Non-gaussian stochastic dynamical model for the el niño southern oscillation. *Physical Review Research*, *4*, L022065. doi: 10.1103/PhysRevResearch.4.L022065

Halpert, M. S., & Ropelewski, C. F. (1992). Surface temperature patterns associated with the southern oscillation. *Journal of Climate*, *5*, 577-593. Retrieved from http://www.jstor.org/stable/26197082

Hamlet, A. F., & Lettenmaier, D. P. (2007, 6). Effects of 20th century warming and climate variability on flood risk in the western u.s. *Water Resources Research*, *43*. doi: 10.1029/2006WR005099

Hanley, D. E., Bourassa, M. A., O'Brien, J. J., Smith, S. R., & Spade, E. R. (2003, 4). A quantitative evaluation of enso indices. *Journal of Climate*, *16*, 1249-1258. doi: 10.1175/1520-0442(2003)16⟨1249:AQEOEI⟩2.0.CO;2

Hastie, T., Tibshirani, R., & Friedman, J. (2009). *The elements of statistical learning.* Springer New York. doi: 10.1007/978-0-387-84858-7

Hempel, S., Frieler, K., Warszawski, L., Schewe, J., & Piontek, F. (2013). A trend-preserving bias correction – the isi-mip approach. *Earth System Dynamics*, *4*, 219-236. doi: 10.5194/esd-4-219-2013

Hurrell, J. W., Holland, M. M., Gent, P. R., Ghan, S., Kay, J. E., Kushner, P. J., ... Marshall, S. (2013, 9). The community earth system model: A framework for collaborative research. *Bulletin of the American Meteorological Society*, *94*, 1339-1360. doi: 10.1175/BAMS-D-12-00121.1

Iqbal, A., & Hassan, S. A. (2018, 4). Enso and iod analysis on the occurrence of floods in pakistan. *Natural Hazards*, *91*, 879-890. doi: 10.1007/s11069-017-3158-y

IRI, C. C. S. (2022). *Enso prediction models.* Retrieved 2022-10-26, from https://iri.columbia.edu/our-expertise/climate/enso/enso-prediction-models/

Isla, F. I., & Junior, E. E. T. (2013, 12). Enso impacts on atlantic watersheds of south america. *Quaternary and Environmental Geosciences*, *4*. doi: 10.5380/abequa.v4i1-2.33032

Kay, J. E., Deser, C., Phillips, A., Mai, A., Hannay, C., Strand, G., ... Vertenstein, M. (2015, 8). The community earth system model (cesm) large ensemble project: A community resource for studying climate change in the presence of internal climate variability. *Bulletin of the American Meteorological Society*, *96*, 1333-1349. doi: 10.1175/BAMS-D-13-00255.1

Kiem, A. S., Franks, S. W., & Kuczera, G. (2003, 1). Multi-decadal variability of flood risk. *Geophysical Research Letters*, *30*. doi: 10.1029/2002GL015992

Kim, H.-M., P. J. W., & Curry, J. A. (2009, 7). Impact of shifting patterns of pacific ocean warming on north atlantic tropical cyclones. *Science*, *325*, 77-80. doi: 10.1126/science.1174062

Kumar, K. K., Rajagopalan, B., Hoerling, M., Bates, G., & Cane, M. (2006, 10). Unraveling the mystery of indian monsoon failure during el niño. *Science*, *314*, 115-119. doi: 10.1126/science.1131152

Kummu, M., Taka, M., & Guillaume, J. H. A. (2018). Gridded global datasets for gross domestic product and human development index over 1990–2015. *Scientific Data*, *5*, 180004. doi: 10.1038/sdata.2018.4

Kundzewicz, Szwed, & Pińskwar. (2019). Climate variability and floods—a global review. *Water*, *11*, 1399. doi: 10.3390/w11071399

Lawrence, D. M., Oleson, K. W., Flanner, M. G., Thornton, P. E., Swenson, S. C., Lawrence, P. J., ... Slater, A. G. (2011, 1). Parameterization improvements and functional and structural advances in version 4 of the community land model. *Journal of Advances in Modeling Earth Systems*, *3*, n/a-n/a. doi: 10.1029/2011MS00045

Lehner, B., & Grill, G. (2013, 7). Global river hydrography and network routing: baseline data and new approaches to study the world's large river systems. *Hydrological Processes*, *27*, 2171-2186. doi: 10.1002/hyp.9740

Liu, Z., & Alexander, M. (2007, 6). Atmospheric bridge, oceanic tunnel, and global climatic teleconnections. *Reviews of Geophysics*, *45*, RG2005. doi: 10.1029/2005RG000172

Lovejoy, S., Amador, L. D. R., & Hébert, R. (2018). *Harnessing butterflies: Theory and practice of the stochastic seasonal to interannual prediction system (stocsips).* Springer International Publishing. Retrieved from `http://link.springer.com/10.1007/978-3-319-58895-7_17` doi: 10.1007/978-3-319-58895-7_17

Lovejoy, S., & Schertzer, D. (2012). *Low-frequency weather and the emergence of the climate.* Retrieved from `http://www.agu.org/books/gm/v196/2011GM001087/2011GM001087.shtml` doi: 10.1029/2011GM001087

Lovejoy, S., & Schertzer, D. (2013). *The weather and climate: Emergent laws and multifractal cascades.* Cambridge University Press. Retrieved from `http://ebooks.cambridge.org/ref/id/CBO9781139093811` doi: 10.1017/CBO9781139093811

L'Heureux, M. L., Tippett, M. K., Takahashi, K., Barnston, A. G., Becker, E. J., Bell, G. D., ... Wang, W. (2019, 2). Strength outlooks for the el niño–southern oscillation. *Weather and Forecasting*, *34*, 165-175. doi: 10.1175/WAF-D-18-0126.1

McCulloch, W. S., & Pitts, W. (1943, 12). A logical calculus of the ideas immanent in nervous activity. *The Bulletin of Mathematical Biophysics*, *5*, 115-133. doi: 10.1007/BF02478259

Neale, R. B., & Group, C. . M. (2012). *Description of the ncar community atmosphere model (cam 5.0) (ncar technical note ncar/tn-486+str)* [report]. Retrieved from `https://www.cesm.ucar.edu/models/cesm1.0/cam/docs/description/cam5_desc.pdf`

Nicholson, S. E., & Kim, J. (1997, 2). The relationship of the el niÑo–southern oscillation to african rainfall. *International Journal of Climatology*, *17*, 117-135. doi: 10.1002/(SICI)1097-0088(199702)17:2⟨117::AID-JOC84⟩3.0.CO;2-O

Nobre, G. G., Jongman, B., Aerts, J., & Ward, P. J. (2017, 8). The role of climate variability in extreme floods in europe. *Environmental Research Letters*, *12*, 084012. doi: 10.1088/1748-9326/aa7c22

Patten, J. M., Smith, S. R., & O'Brien, J. J. (2003). Impacts of enso on snowfall frequencies in the united states. *Weather and Forecasting*, *18*, 965-980. doi: 10.1175/1520-0434(2003)018⟨0965:IOEOSF⟩2.0.CO;2

Penland, C., & Magorian, T. (1993, 6). Prediction of niño 3 sea surface temperatures using linear inverse modeling. *Journal of Climate*, *6*, 1067-1076. doi: 10.1175/1520-0442(1993)006⟨1067:PONSST⟩2.0.CO;2

Penland, C., & Sardeshmukh, P. D. (1995, 8). The optimal growth of tropical sea surface temperature anomalies. *Journal of Climate*, *8*, 1999-2024. doi: 10.1175/1520-0442(1995)008⟨1999:TOGOTS⟩2.0.CO;2

Pielke, R. A., & Landsea, C. N. (1999, 10). La niña, el niño and atlantic hurricane damages in the united states. *Bulletin of the American Meteorological Society*, *80*, 2027-2033. doi: 10.1175/1520-0477(1999)080⟨2027:LNAENO⟩2.0.CO;2

Rashid, H. A. (2020, 12). Factors affecting enso predictability in a linear empirical model of tropical air-sea interactions. *Scientific Reports*, *10*, 3931. doi: 10.1038/s41598-020-60371-1

Ropelewski, C. F., & Halpert, M. S. (1986, 12). North american precipitation and temperature patterns associated with the el niño/southern oscillation (enso). *Monthly Weather Review*, *114*, 2352-2362. doi: 10.1175/

1520-0493(1986)114⟨2352:NAPATP⟩2.0.CO;2
Saghafian, B., Haghnegahdar, A., & Dehghani, M. (2017, 5). Effect of enso on annual maximum floods and volume over threshold in the southwestern region of iran. *Hydrological Sciences Journal*, *62*, 1039-1049. doi: 10.1080/02626667.2017.1296229
Shi, W. (2007). *Global daily surface air temperature analyses* [dataset]. Retrieved from https://ftp.cpc.ncep.noaa.gov/precip/PEOPLE/wd52ws/global_temp/CPC-GLOBAL-T.pdf
Shukla, J., & Paolino, D. A. (1983, 9). The southern oscillation and long-range forecasting of the summer monsoon rainfall over india. *Monthly Weather Review*, *111*, 1830-1837. doi: 10.1175/1520-0493(1983)111⟨1830:TSOALR⟩2.0.CO;2
Tan, X., Tang, Y., Lian, T., Yao, Z., Li, X., & Chen, D. (2020). A study of the effects of westerly wind bursts on enso based on cesm. *Climate Dynamics*, *54*, 885-899. Retrieved from https://doi.org/10.1007/s00382-019-05034-2 doi: 10.1007/s00382-019-05034-2
Tsonis, A. A. (2018). *Insights in climate dynamics from climate networks.* Springer International Publishing. Retrieved from http://link.springer.com/10.1007/978-3-319-58895-7_29 doi: 10.1007/978-3-319-58895-7_29
Tsonis, A. A., Swanson, K. L., & Roebber, P. J. (2006, 5). What do networks have to do with climate? *Bulletin of the American Meteorological Society*, *87*, 585-596. Retrieved from http://journals.ametsoc.org/doi/10.1175/BAMS-87-5-585 doi: 10.1175/BAMS-87-5-585
Vega-Westhoff, B., & Sriver, R. L. (2017, 12). Analysis of enso's response to unforced variability and anthropogenic forcing using cesm. *Scientific Reports*, *7*, 18047. doi: 10.1038/s41598-017-18459-8
Wang, C. (2018, 11). A review of enso theories. *National Science Review*, *5*, 813-825. doi: 10.1093/nsr/nwy104
Wang, H.-J., Zhang, R.-H., Cole, J., & Chavez, F. (1999). El niño and the related phenomenon southern oscillation (enso): The largest signal in interannual climate variation. *Proceedings of the National Academy of Sciences*, *96*, 11071-11072. doi: 10.1073/pnas.96.20.11071
Wang, X., Slawinska, J., & Giannakis, D. (2020, 12). Extended-range statistical enso prediction through operator-theoretic techniques for nonlinear dynamics. *Scientific Reports*, *10*, 2636. doi: 10.1038/s41598-020-59128-7
Ward, P., Kummu, M., & Lall, U. (2016, 8). Flood frequencies and durations and their response to el niño southern oscillation: Global analysis. *Journal of Hydrology*, *539*, 358-378. doi: 10.1016/j.jhydrol.2016.05.045
Ward, P. J., Eisner, S., Flörke, M., Dettinger, M. D., & Kummu, M. (2014). Annual flood sensitivities to el niño–southern oscillation at the global scale. *Hydrology and Earth System Sciences*, *18*, 47-66. doi: 10.5194/hess-18-47-2014
Ward, P. J., Jongman, B., Kummu, M., Dettinger, M. D., Sperna Weiland, F. C., & Winsemius, H. C. (2014). Strong influence of el niño southern oscillation on flood risk around the world. *Proceedings of the National Academy of Sciences*, *111*, 15659-15664. doi: 10.1073/pnas.1409822111
Ward, P. J., Jongman, B., Salamon, P., Simpson, A., Bates, P., Groeve, T. D., ... Winsemius, H. C. (2015, 8). Usefulness and limitations of global flood risk models. *Nature Climate Change*, *5*, 712-715. doi: 10.1038/nclimate2742
Weng, H., Behera, S. K., & Yamagata, T. (2009, 4). Anomalous winter climate conditions in the pacific rim during recent el niño modoki and el niño events. *Climate Dynamics*, *32*, 663-674. doi: 10.1007/s00382-008-0394-6
Winsemius, H. C., Beek, L. P. H. V., Jongman, B., Ward, P. J., & Bouwman, A. (2013, 5). A framework for global river flood risk assessments. *Hydrology and Earth System Sciences*, *17*, 1871-1892. doi: 10.5194/hess-17-1871-2013
Xie, P., Chen, M., Yang, S., Yatagai, A., Hayasaka, T., Fukushima, Y., & Liu, C. (2007, 6). A gauge-based analysis of daily precipitation over east asia. *Journal*

of Hydrometeorology, *8*, 607-626. doi: 10.1175/JHM583.1

Yamazaki, D., Kanae, S., Kim, H., & Oki, T. (2011, 4). A physically based description of floodplain inundation dynamics in a global river routing model. *Water Resources Research*, *47*. doi: 10.1029/2010WR009726

Yan, Y., Wu, H., Gu, G., Ward, P. J., Luo, L., Li, X., … Tao, J. (2020, 11). Exploring the enso impact on basin-scale floods using hydrological simulations and trmm precipitation. *Geophysical Research Letters*, *47*. doi: 10.1029/2020GL089476

Yang, S., Li, Z., Yu, J.-Y., Hu, X., Dong, W., & He, S. (2018, 11). El niño–southern oscillation and its impact in the changing climate. *National Science Review*, *5*, 840-857. doi: 10.1093/nsr/nwy046

Yu, J.-Y., & Zou, Y. (2013, 3). The enhanced drying effect of central-pacific el niño on us winter. *Environmental Research Letters*, *8*, 014019. doi: 10.1088/1748-9326/8/1/014019

Zhang, W., Leung, Y., & Fraedrich, K. (2015, 6). Different el niño types and intense typhoons in the western north pacific. *Climate Dynamics*, *44*, 2965-2977. doi: 10.1007/s00382-014-2446-4

Zhou, D., Gozolchiani, A., Ashkenazy, Y., & Havlin, S. (2015, Dec). Teleconnection paths via climate network direct link detection. *Phys. Rev. Lett.*, *115*, 268501. Retrieved from `https://link.aps.org/doi/10.1103/PhysRevLett.115.268501` doi: 10.1103/PhysRevLett.115.268501